ELSEVIER

Critical Reviews in Oncology/Hematology 65 (2008) 32–42

CRITICAL REVIEWS IN
Oncology
Hematology
Incorporating Geriatric Oncology

www.elsevier.com/locate/critrevonc

Promoting evidence-based management of anemia in cancer patients: Background, development, and scientific validation of RESPOND, a web-based clinical guidance system based on the EORTC guidelines

Matti Aapro[a], Joanna Van Erps[b], Karen MacDonald[c], Pierre Soubeyran[d,e], Matthew Turner[f], Michael Muenzberg[g], Robert Dunlop[h], Hans Warrinnier[i], Ivo Abraham[c,j,k,l,*]

[a] *Institut Multidisciplinaire d'Oncology, Clinique de Genolier, 1 Route du Muids, CH-1272 Genolier, Switzerland*
[b] *Afdeling Oncologie en Hematologie, Algemeen Stedelijk Ziekenhuis Aalst, Merestraat 80, B-9300 Aalst, Belgium*
[c] *Matrix45, 620 Frays Ridge Road, Earlysville, VA 22936, USA*
[d] *Institut Bergonié, 229, Cours de l'Argonne, F-33076 Bordeaux, France*
[e] *Université Victor Segalen Bordeaux 2, 146 Rue Léo Saignat, F-33076 Bordeaux, France*
[f] *Roche Products Limited, 6 Falcon Way, Shire Park, Welwyn Garden City, Hertfordshire AL7 1TW, United Kingdom*
[g] *F. Hoffmann-La Roche AG, Grenzacherstrasse, Bau 74, CH-4070 Basel, Switzerland*
[h] *InferMed, 25 Bedford Square, London WC1B 3HW, United Kingdom*
[i] *Roche, Dantestraat 75, B-1070 Brussel, Belgium*
[j] *Center for Health Outcomes and Policy Research, University of Pennsylvania, 418 Curie Boulevard, Philadelphia, PA 19104, USA*
[k] *Leonard Davis Institute of Health Economics, University of Pennsylvania, 418 Curie Boulevard, Philadelphia, PA 19104, USA*
[l] *Institute of Aging, University of Pennsylvania, 418 Curie Boulevard, Philadelphia, PA 19104, USA*

Accepted 2 May 2007

Contents

1. Introduction 33
2. Background 33
   2.1. Management of cancer-related anemia: from evidence to guidelines to practice 33
   2.2. Increasing the congruence of evidence-based guidelines and actual clinical practice through clinical guidance systems 34
3. The RESPOND validation methodology 34
   3.1. Concept 34
   3.2. Study populations 35
      3.2.1. Overview 35
      3.2.2. Patient selection criteria for Step 2 35
   3.3. Step 1: Reliability and content validity 35
      3.3.1. Design 35
      3.3.2. Algorithm evaluation and calculation of the intraclass correlation coefficient 35
   3.4. Step 2: Concurrent and discriminant validity 35
      3.4.1. Design, tool, and samples 35
      3.4.2. Data model 38
      3.4.3. Calculation of the congruence score 38
      3.4.4. Statistical considerations 38
4. Discussion 39
   Disclosures 40
      Section I: Contributions to the project 40
      Section II: Conflict of interest 40
      Section III: Sponsor and role of sponsor 40

* Corresponding author at: 620 Frays Ridge Road, Earlysville, VA 22936, USA. Tel.: +1 303 997 2697; fax: +1 978 945 8374.
E-mail address: iabraham@matrix45.com (I. Abraham).

1040-8428/$ – see front matter
doi:10.1016/j.critrevonc.2007.05.002

Section IV: Role of medical writer or editor 41
Reviewers 41
Acknowledgments 41
References 41
Biographies 41

Abstract

The 2006 EORTC guidelines for erythropoietic proteins in cancer-related anemia provide the most up-to-date assessment of the evidence base. Considering general concerns in medicine about clinicians' adoption of evidence-based guidelines, it is critical to find ways of bringing guidelines to the point of care. We describe the rationale behind RESPOND, a web-based clinical guidance system based on the EORTC guidelines, and the methodologies of two studies conducted to validate the system. In a first descriptive study, experts are asked to rate the accuracy of every algorithm derived from the guidelines. In a second step and using a hybrid matched pre-post design, separate retrospective and prospective patient cohorts matched by type of cancer and similarity of chemotherapy regimen (33 pairs) are used to examine the extent to which clinicians' practice patterns converge with the EORTC guidelines when they use or not use the RESPOND system. It is hypothesized that these studies will provide the necessary validation for RESPOND as an evidence-based clinical support tool.

Keywords: Erythropoietic proteins; Anemia; Cancer; Treatment guidelines; Decision-support systems; Validation

1. Introduction

Anemia is one of the most prevalent, major side effects of cancer and cancer treatment [1,2]. Attributed to such factors as inflammation-induced changes in iron homeostasis, production of erythroid progenitor cells, erythropoietin production, and the life span of red blood cells [3] as well as tumor hypoxia [4], cancer-related anemia is associated with poor prognosis and treatment outcomes [4–6] and has a significant impact on patients' quality of life [4,7]. In the European Cancer Anaemia Survey (ECAS) [2], almost 70% of cancer patients experienced anemia (hemoglobin (Hb) < 12g/dL) during a 6-month period, yet only 39.8% of them were treated: 17.4% received an erythropoiesis stimulating agent (ESA) (either alone or in combination with blood transfusion or supplementary iron), 14.9% were given blood transfusion alone, and 6.5% were treated with iron alone.

Given the significant therapeutic benefits of ESAs [8–10], but also the variability in treatment patterns and outcomes [1,2], evidence-based guidelines have been proposed by the American Society of Clinical Oncology (ASCO) and the American Society of Hematology (ASH) [8] and, more recently, by the European Society for Research and Treatment of Cancer (EORTC) [9,10]. These guidelines provide evidence-based direction to clinicians in the (differential) diagnosis and management of anemia in cancer patients.

However, there is significant concern in medicine at large about the extent to which evidence-based guidelines lead to *de facto,* evidence-based decision-making about individual patient care [11]. Knowledge about and adoption of guidelines by clinicians has been judged – intuitively rather than empirically – to be mixed and disappointing [12], and has led to a call for developing strategies, methods, and tools to bring evidence closer to the point of care.

In this article, we describe the rationale and background behind the RESPOND system, as well as the methodology used in a comprehensive system validation effort. RESPOND is a web-based clinical guidance system for the (differential) diagnosis and management of anemia in patients with cancer based on the 2006 Update of the EORTC guidelines [10] and intended for point-of-care use by physicians, nurses, and other cancer clinicians.

2. Background

2.1. Management of cancer-related anemia: from evidence to guidelines to practice

Erythropoietin is a naturally occurring glycoprotein that is synthesized mainly in the kidneys and induces proliferation and differentiation of precursor stem cells into mature red blood cells. The human erythropoietin gene was cloned in 1985, and several forms of the recombinant hormone (rHuEPO) are now available for clinical use. These erythropoiesis stimulating agents (ESAs) are potent stimulators of early erythroid progenitor cells in the bone marrow [13,14].

Treatment with ESAs to stimulate erythropoiesis and to increase Hb levels has been found to be an efficacious and safe way to manage cancer-related anemia [8–10] and to improve patients' function and quality of life [4]. Despite the demonstrated benefits of ESAs in controlled studies, the "real-world" management of cancer-related anemia with ESAs, as documented, for instance, in the 2001 data from the ECAS study [2], is suboptimal and characterized by considerable variability in the practice patterns of prescribing physicians.

By setting clinical targets and providing practice recommendations, the ASCO/ASH [8] and EORTC [9,10] guidelines have the potential to decrease this variability in practice patterns and improve patient outcomes. However, the extent to which physicians are familiar with these guidelines

and have adopted them into their clinical practice is unknown. If other conditions are any indication, this adoption should not be assumed.

2.2. Increasing the congruence of evidence-based guidelines and actual clinical practice through clinical guidance systems

One method to increase the congruence of clinical practice with guidelines is to design computerized systems that provide clinicians with evidence-based guidance based on patient data entered [15–17]. Many decision support systems for use in clinical practice have been developed [18,19], but few have been tested in randomized controlled trials (RCTs). The most effective systems are those that are integrated into the workflow, provide guidance at the point of care, and provide actionable recommendations [20]. In fact, several systems have not had measurable impacts on patient outcomes [20], perhaps because of inadequacies in such factors as development, design, user interface, validation, acceptance by clinicians, and integration at the point of care. Moreover, most of these systems are not evidence-based but knowledge-based and try to emulate and formalize processes of clinical inference [17].

In an initial effort to make the ASH/ASCO guidelines available at the point of care a prototype web-based system (ACT, Anemia Control and Testing) was developed in 2003–2004 under the auspices of Roche-Belgium. In this prototype, Hb (measured by a point-of-care testing device [HemoCue, Ängelholm, Sweden]) and other clinical parameters were analyzed through a set of algorithms derived from the ASH/ASCO guidelines. Recommendations based upon these guidelines were then generated. With the publication of the EORTC guidelines in 2004 [9], F. Hoffmann-La Roche AG (Basel, Switzerland) proposed to adopt the prototype's core concept; to support the design, development, and validation of RESPOND, a web-based clinical guidance system based on the 2006 Update of the EORTC guidelines [10], and to make this system available to cancer clinicians in Europe, if not worldwide.

From the onset, cancer anemia experts (Aapro, Soubeyran, Van Erps), evidence-based healthcare researchers (Abraham, MacDonald), and sponsors (Turner, Muenzberg, Warrinnier) emphasized the need for rigorous testing and validation of the RESPOND system, including assessment of validity and reliability—to use psychometric terms [21]. Remarkably, while there is extensive literature on clinical software validation we have found no publications providing empirical evidence for the validity of clinical decision support systems driven by evidence-based guidelines. It seems as if the underlying assumption is that "if the algorithms could be programmed, then the system must be adequate".

Guidelines like those proposed by the EORTC try to be as complete and exhaustive as possible. However, the fact that some evidence may be of lower grades of strength, or absent altogether, makes that even the best guidelines may have incomplete or (as of yet) inconclusive recommendations. Consequently, while many guidelines may translate rather easily into programmable algorithms, others require additional interpretation and operationalization before they can be formalized into algorithms. It is in this latter process that the risk for incorrect analysis, erroneous inference, and inappropriate clinical guidance is the highest.

We have adopted a critical, if not reluctant, perspective on computerized systems that enable clinicians to make evidence-based clinical decisions at the point of care. Intrinsically, we are committed to the design of RESPOND as a web-based decision support tool to provide clinical guidance to cancer clinicians in the management of cancer-related anemia. On the other hand, we oppose having the guidelines programmed into a computer system without verification and testing. Therefore, and drawing upon principles of psychometric theory [21], the validation of RESPOND is comprised of two steps. First, a group of experts must confirm that the algorithms correctly translate the EORTC guidelines into clinical recommendations. Second, RESPOND must be tested in a clinical setting to examine the extent to which its use is associated with clinical practice that is congruent with the EORTC guidelines.

3. The RESPOND validation methodology

3.1. Concept

The general aim of this validation study is to assess various dimensions of the validity of RESPOND. Thus this study aims to first investigate whether RESPOND correctly translates the EORTC guidelines into clinical recommendations. Once this has been established, the study then intends to investigate whether the use of RESPOND is associated with practice patterns aligned with the EORTC guidelines.

The flowchart in Fig. 1 summarizes the process of this study:

1. *Content validity and accuracy of the algorithms*. Are the algorithms embedded in RESPOND accurate operationalizations of the assessment and management rules specified by or embedded in the EORTC guidelines?
2. *Concurrent validity of RESPOND*. To what extent are clinicians' anemia care practice patterns congruent with EORTC guidelines when they are using the patient monitoring and decision-support functionality of RESPOND?
3. *Discriminant validity of RESPOND*. Are the anemia care practice patterns of clinicians using the RESPOND system more congruent with EORTC guidelines than practice patterns generated in the absence of the RESPOND system? In other words, is using the RESPOND system associated with practice patterns that are more congruent with the EORTC guidelines when compared to practice patterns observed in the absence of the RESPOND system?

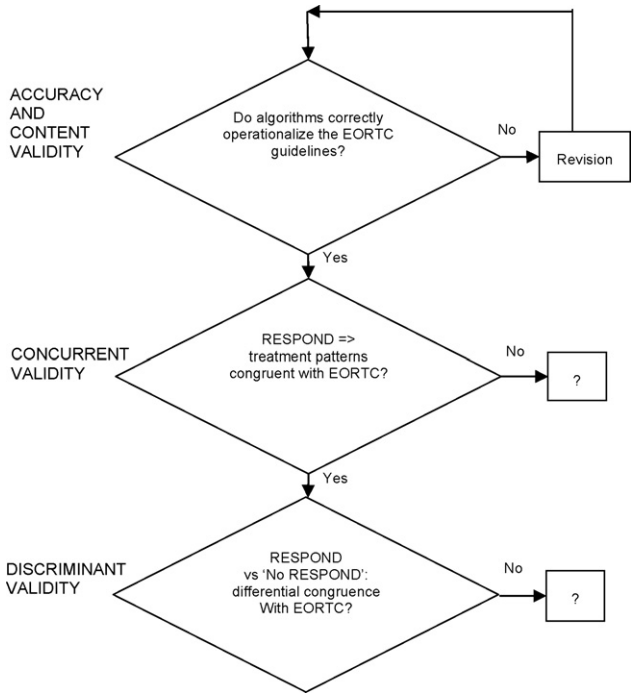

Fig. 1. Flowchart of the study objectives.

3.2. Study populations

3.2.1. Overview

For Step 1: physicians and nurses with an established scientific and clinical track record in the area of cancer-related anemia; including members of the EORTC task force that developed the 2006 revised guidelines and experts in evidence-based healthcare.

For Step 2: adult cancer patients with anemia and undergoing chemotherapy matched by type of cancer and chemotherapy regimen. Any treatment for cancer, anemia, or other clinical condition is per the treating physician's best clinical judgment.

There are no required treatment protocols for this study. All prescribing must be in accordance with the approved label/scientific leaflet in Belgium.

3.2.2. Patient selection criteria for Step 2

Patients must be age 18 or older, diagnosed with a solid or hematological malignancy, treated with chemotherapy, and anemic (Hb $\leq$ 11.0 g/dL). Patients in the post-cohort must provide written informed consent. Excluded are patients who have undergone autologous or allogeneic blood stem cell transplantation, or who are diagnosed with myelodysplastic syndrome.

3.3. Step 1: Reliability and content validity

3.3.1. Design

The first step of this study employs a descriptive design. A purposively selected panel of experts is asked to rate the accuracy of the EORTC-based algorithms developed for the RESPOND system. It is assumed that if a given algorithm is judged to be accurate, RESPOND will not only feature a reliable analysis model for a given guideline but also adequately cover the content of that guideline. In the aggregate, across guidelines, it will permit the conclusion that the RESPOND algorithms reliably and validly represent the EORTC guidelines.

3.3.2. Algorithm evaluation and calculation of the intraclass correlation coefficient

Experts are presented by email with a document listing each EORTC guideline, the algorithm(s) operationalizing a given guideline, the required data elements, and the guideline recommendation(s) associated with each outcome of a given algorithm (see Table 1 for an example). For each algorithm, experts are asked to indicate with a binary score (yes/no) whether the algorithm is an accurate representation of the referent EORTC guideline. For any "no" answer, the expert was asked to explain why and to offer an alternate formulation.

The intraclass correlation coefficient (ICC) [22] is calculated for each algorithm. If the ICC > 0.90, the algorithm will be considered accepted. Any algorithm with ICC $\leq$ 0.90 is revised per the suggestions of the dissenting experts and resubmitted by email to the panel until the ICC > 0.90.

In the event that after 3 email voting rounds a given algorithm's ICC is still $\leq$0.90, the algorithm is discussed in a conference call of the experts in order to achieve consensus of at least 90% of the experts.

3.4. Step 2: Concurrent and discriminant validity

Once the reliability and content validity have been established, it is important to examine (1) whether clinicians' practice patterns are congruent with EORTC guidelines when they use RESPOND (concurrent validity) and (2) whether clinicians' practice patterns are congruent with EORTC guidelines when using RESPOND versus not using RESPOND (discriminant validity).

Especially the comparison of practice patterns when using versus not using RESPOND is methodologically challenging. Note that, for exploratory purposes, the study also examines associations between the use of RESPOND and clinical patient characteristics, and between congruence scores and clinical patient characteristics.

This part of the study is being done in a mixed oncology-hematology center in a non-university teaching hospital in Belgium. A one-center design suffices for the assessment of concurrent and discriminant validity. Note that the overall study protocol (steps 1 and 2) was approved by the Ethical Committee of the Algemeen Stedelijk Ziekenhuis Aalst (Aalst, Belgium), where the study is being conducted.

3.4.1. Design, tool, and samples

The second step employs a hybrid matched pre-post design, in which separate pre (retrospective) and post

Table 1
Examples of selected EORTC guidelines, proposed algorithm, and expert voting procedures

Guideline 2 Recommendation.

In cancer patients receiving chemotherapy and/or radiotherapy, treatment with erythropoietic proteins should be initiated at a Hb level of 9-11 g/dL based on anaemia-related symptoms (grade A).

Set #	Criteria (if met, then guideline recommendation provided)	Accuracy (Y/N)	If no, then please indicate your suggestions for revision.
2.1	((if chemo=yes) AND/OR (if radio=yes)) AND (if auto transplant=no) AND (if allo transplant=no) AND ((if Hb > 8.9 g/dL) AND (if Hb < 11.1 g/dL)) AND (if symptoms=yes) AND (if epo=no)	N ☐ → Y ☐ ↓	

Guideline 5 Recommendation.

Erythropoietic proteins may be considered in asymptomatic, anaemic patients with a Hb level of ≤ 11.9 g/dL to prevent a further decline in Hb, according to individual factors (e.g., type/intensity of chemotherapy, baseline Hb) and the duration and type of further planned treatment (grade B).

Set #	Criteria (if met, then guideline recommendation provided)	Accuracy (Y/N)	If no, then please indicate your suggestions for revision.
5.1	(if auto transplant=no) AND (if allo transplant=no) AND ((if Hb > 8.9* g/dL) AND (if Hb < 12.0 g/dL)) AND (if symptoms=no) AND (if epo=no)	N ☐ → Y ☐ ↓	

* Note: lower Hb limit included in algorithm since Hb < 9.0 g/dL would trigger guideline recommendation #4.

Guideline 6 Recommendation.

We do not recommend the prophylactic use of erythropoietic proteins to prevent anaemia in patients undergoing chemotherapy and/or radiotherapy who have normal Hb values at the start of treatment (grade A).

Set #	Criteria (if met, then guideline recommendation provided)	Accuracy (Y/N)	If no, then please indicate your suggestions for revision.
6.1	((if chemo=yes) OR (if radio=yes)) AND (if auto transplant=no) AND (if allo transplant=no) AND (if Hb >11.9 g/dL) AND (if gender = F)	N ☐ → Y ☐ ↓	
6.2	((if chemo=yes) OR (if radio=yes)) AND (if auto transplant=no) AND (if allo transplant=no) AND (if Hb >13.9 g/dL) AND (if gender = M)	N ☐ → Y ☐ ↓	

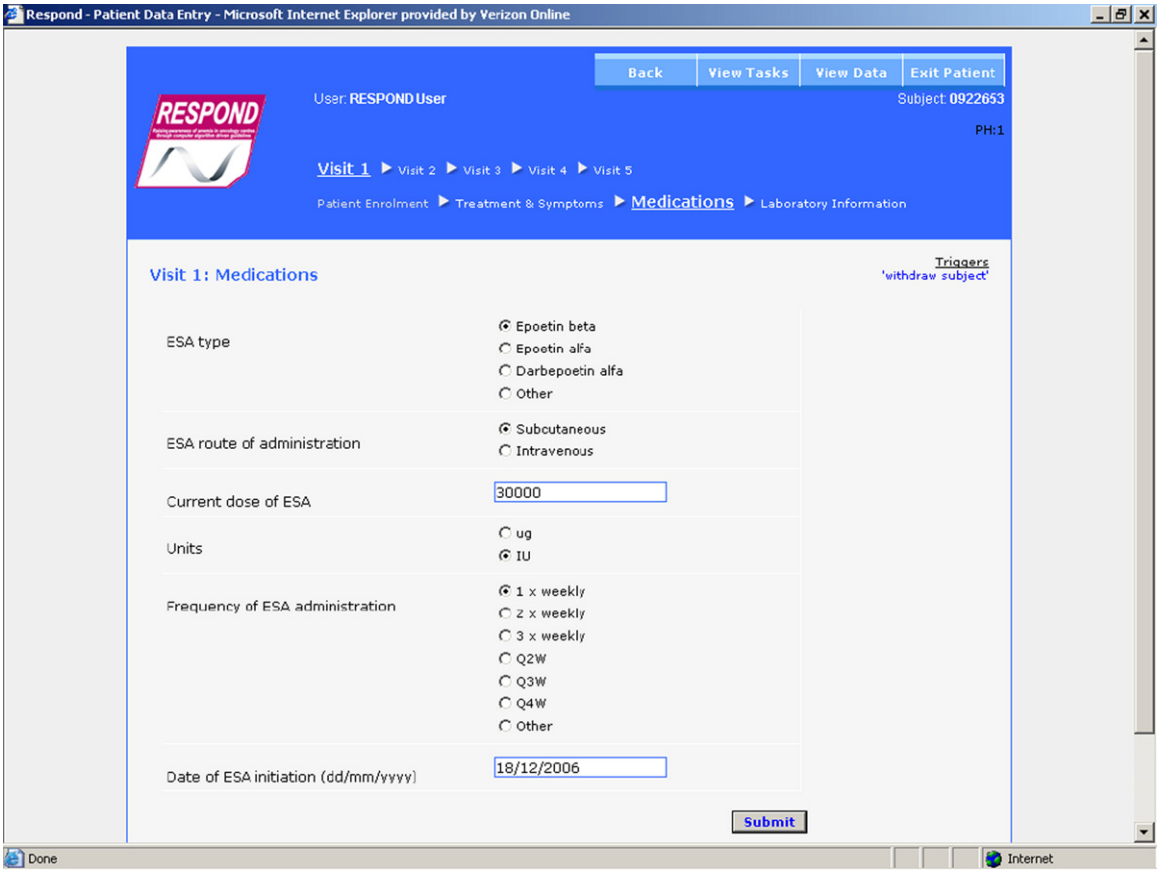

Fig. 2. Example of a data entry screen in RESPOND.

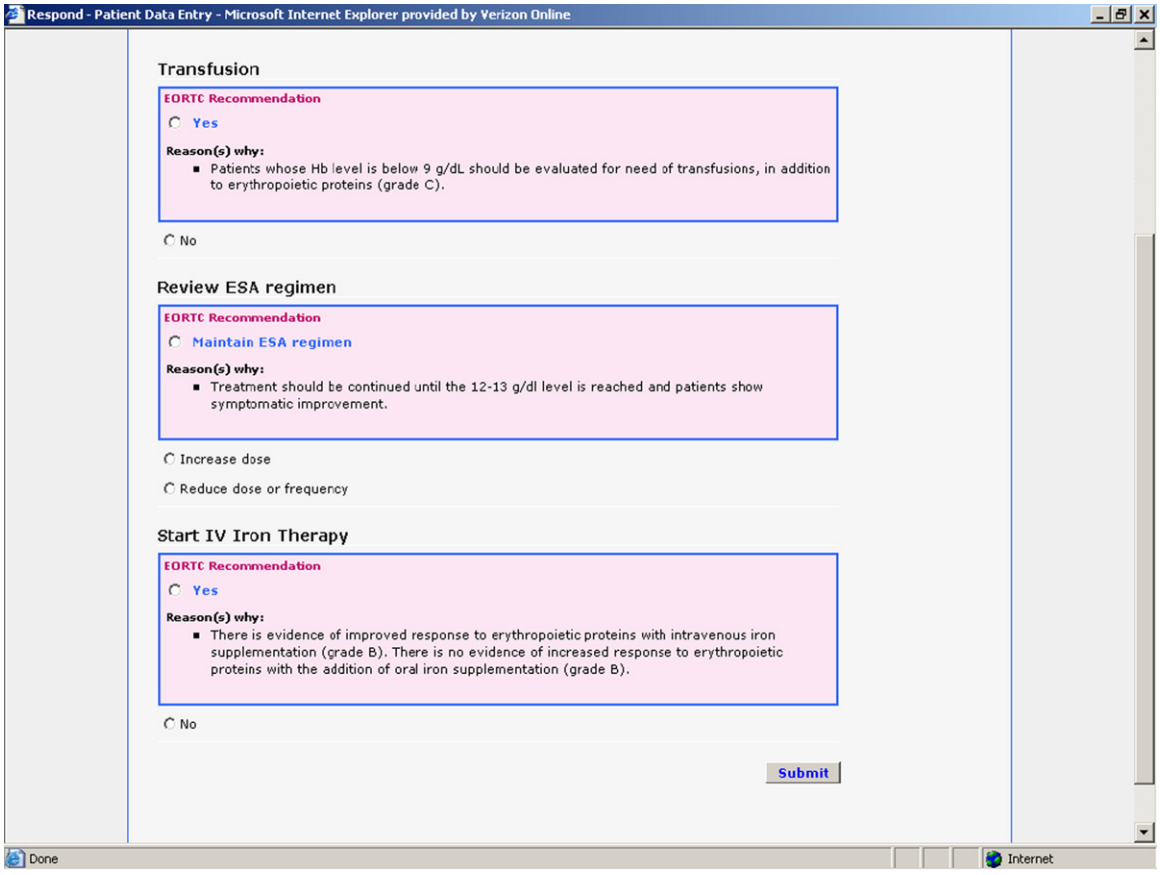

Fig. 3. Example of a clinical guidance screen in RESPOND.

(prospective) matched cohorts are established. Patients are matched by type of cancer and similarity of chemotherapy regimen.

Clinicians are asked to use the system prospectively in the care of a sample of 33 consenting adult patients being treated for cancer with anemia (see Figs. 2 and 3 for examples of input and recommendation screens). This sample constitutes the prospective post-cohort ("post" referring to "after" introduction of the RESPOND tool).

Concurrent with the prospective component, staff review the records of anemic cancer patients who underwent chemotherapy at the center between 4 and 12 months prior to the introduction of the RESPOND tool and for whom data from at least 4 consecutive visits over 4 months are available. For each patient recruited into the post-cohort, center staff identifies potential patients with the same type of cancer and similar chemotherapy regimen. This information is sent to an independent investigator for matching. A total of 33 patients will thus be selected to constitute the pre-cohort ("pre" referring to "prior to" introduction of the RESPOND tool).

For each patient in the study (i.e., both pre- and post-cohort patients), a congruence score (CS) is calculated. This score quantifies the extent to which the care provided to a given patient is in accordance with the EORTC guidelines. The scores of the post-cohort will be used to assess concurrent validity. The scores of the matched pre- and post-cohorts will be used to determine discriminant validity. The following schematic summarizes Step 2:

2 matched samples (N=33 each): sample 1: 4 months prospective data after RESPOND (post) sample 2: 4 months retrospective data prior to RESPOND (pre)	Concurrent validity: Congruence scores of sample 1 (post) Discriminant validity: Difference between sample 2 (pre) and sample 1 (post) congruence scores

Several other design options were considered but excluded. Randomization of patients within the same center to RESPOND-driven versus standard care is likely to be confounded by within-clinician spill-over effects of knowledge, experience, and perspective. Randomization of clinicians or centers to RESPOND-driven versus standard care is likely to be confounded by resentment and compensation among those clinicians and centers randomized to standard care. This is also likely to be the case in a clinician or center cross-over design for those randomized to start with standard care. Note that standard care does not preclude knowledge and application of EORTC guidelines, as these have been widely disseminated.

Table 2
Calculation of the congruence score

Item	Scoring
Rule out/treat other causes of anemia: iron deficiency	+0.25
Rule out/treat other causes of anemia: bleeding	+0.25
Rule out/treat other causes of anemia: nutritional deficits	+0.25
Rule out/treat other causes of anemia: hemolysis	+0.25
ESA treatment initiated at Hb range 9.0–11.0 g/dL	+1.00
If Hb < 9.0 g/dL patient was considered for blood transfusion; OR: no need for blood transfusion.	+1.00
Anemic asymptomatic patient with Hb ≤ 11.9 g/dL was considered for ESA treatment on individual basis; OR: patient does not fall in this category	+1.00
Hb target set at Hb range 12.0–13.0 g/dL	+1.00
Initial ESA dose = 40,000 IU epoetin alfa Q1W; 30,000 IU epoetin beta Q1W; or darbepoetin alfa Q1W or Q3W	+1.00
Initial ESA dose was fixed and not weight adjusted	+1.00
Hb target achieved at 4–8 weeks; if not, individualized dose escalation	+1.00
Hb levels maintained in Hb range 12.0–13.0 g/dL for up to 8 weeks (+0.125 per week)	+1.00
ESA discontinued if Hb ≥ 13.0 g/dL (or Hb in range 12.0–13.0 g/dL at end of study period)	+1.00
Maximum possible score	10.00

3.4.2. Data model[1]

As Step 2 is an observational study, only those data routinely collected (or available from the medical record) are entered into the RESPOND system. Enrollment data collected include eligibility data, patient demographics, cancer history, anemia history, and relevant comorbidities. At each of the 4 visits, the following data are collected as available: weight, hematological parameters, serum ferritin, transferrin saturation, serum creatinine, serum albumin, C-reactive protein, ESA treatment, iron supplementation, blood transfusions, cancer treatment, other medications, bleeding episodes, signs and symptoms of anemia, performance status, and signs of malnutrition.

3.4.3. Calculation of the congruence score

The congruence score (CS) quantifies the extent to which a patient's anemia assessment, management, and outcomes are in accordance with the EORTC guidelines. It was developed by a subpanel of the author team and submitted for review and evaluation by all other authors. The CS reflects the 10 key decisions that must be made in the use of ESAs to be in accordance with the EORTC guidelines:

1. Ruling out four other causes of anemia.
2. Initiating ESA treatment at Hb range 9.0–11.0 g/dL.
3. Considering/ruling out the need for blood transfusion.
4. Considering/ruling out the possibility of asymptomatic presentation of anemia.
5. Setting the target Hb range at 12.0–13.0 g/dL.
6. Using the correct initial ESA dose.
7. Using an initial fixed dose, not a weight-adjusted dose.
8. Achieving target at 4–8 weeks; otherwise, initiate an individualized dose escalation.
9. Maintaining Hb in the 12.0–13.0 g/dL range for up to eight weeks.
10. Discontinuing ESA treatment if Hb ≥ 13.0 g/dL.

[1] Complete data model is available from the corresponding author.

For each key decision made in a patient's care, one point is given for range of 0-10 total points. For all but two of the key decisions, the point is given in whole units. However, for the decision regarding other causes of anemia (decision 1), 0.25 point is given for each of the four other causes ruled out; i.e., ruling out all four generates a full point. Likewise, for the guideline recommending maintaining Hb in range for up to 8 weeks, 0.125 point is given for each week that the patient's Hb level is in this range.

The congruence score is calculated for all 66 patients. Table 2 summarizes the CS calculation.

3.4.4. Statistical considerations

The sample size of 33 patients in each cohort (66 total) is sufficient for detecting mean differences of 0.50 standardized units under the non-central *t* distribution and correlation coefficients of 0.45 with power of 0.80 at alpha of 0.05.

Testing the concurrent validity of RESPOND will be done using the one-sample *t*-test to examine the null hypothesis that the CS equals 10, i.e., that all key decisions embedded in the EORTC guidelines have been made.

Testing for statistical significance between the matched pre- and post-cohort will be done using the paired *t*-test (*t*-test for dependent samples) or, in the case of violation of statistical assumptions, its nonparametric analog. This test will examine whether or not the two matched subsamples show statistically significant differences in the CS. This is a paired test and the basic unit of calculation is the standardized difference between the two paired patients on the CS. The null hypothesis to be tested is that the standardized difference approximates zero, in which case the assumption of discriminant validity will be rejected. If across all pairs of matched patients, the standardized difference exceeds a critical value, it will be possible to infer the discriminant validity of RESPOND.

The secondary exploratory analyses on the potential association between the use of RESPOND and clinical

patient characteristics and the association between congruence scores and clinical patient characteristics will use measures of association. Measures will be selected under consideration of the levels of measurement of the variables involved.

4. Discussion

In an effort to develop a well-tested and validated system, the RESPOND initiative merges conventional methods of software design and development with a dual scientific purpose of expert evaluation of content and field testing of the system in actual patient situations. Certainly, the psychometric-like assessment, evident as it may appear to behavioral and social scientists, is rather unusual in health care. Typically, the (implicit) assumption in health care is that, once evidence-based guidelines have passed the test of peer review and have been published in a refereed journal, they are ready to be made available to clinicians through various educational, print, or electronic means—and that improved patient outcomes will follow.

Step 1 has been completed and the findings are reported elsewhere [23]. Summarized, consensus was achieved (ICC=1.00) on all but 5 algorithmic definitions after one round of review. Following revision, the remaining 5 definitions were approved unanimously as well (ICC = 1.00).

The major focus of the Step 2 component, currently in progress, is the assessment of the congruence of clinicians' practice patterns with the EORTC guidelines in and between two subsamples of patients. Admittedly, the impact of RESPOND on patient outcomes is not (yet) a part of the RESPOND initiative and clinicians in the Step 2 process are instructed to practice according to their best clinical judgment, including overriding recommendations generated by RESPOND. Examining the clinical impact of RESPOND is a necessary next step, preferably in the form of a cluster randomized controlled trial or, alternately, a carefully designed observational study. However, we believe it would be ethically questionable to subject patients to an intervention (c.q., diagnosis and management of cancer-related anemia under the guidance of a decision-support tool) if the technology girding the intervention has not been thoroughly tested and validated. Perhaps a stretch, but to draw an analogy with oncological drug development, the work described in this article is comparable to the discovery phase (conceptualization of the system), pre-clinical testing and proof of concept (reliability and content validity assessment), initial formulation (programming), and safety and (clinician) tolerability testing using patients (Step 1).

Part of the need for testing and validation is also safety (and by extension liability) driven. In the end, a system like RESPOND may provide clinical advice that is followed by clinicians and made part of patient's therapeutic plan. As with any intervention (or, for that matter, medical device), the necessary precautions and actions must have been taken to assure that the system is safe—even before it may be found to be clinically effective.

There are several additional challenges that will need to be addressed as RESPOND evolves into a clinically useful support tool. First, the knowledge regarding the use of ESAs in cancer patients with anemia is likely to evolve. Guidelines in place today are likely to change as new evidence about the effectiveness and safety of ESAs in supportive cancer care becomes available. RESPOND, like systems incorporating evidence-based guidelines in other areas of healthcare, will need to be adapted as the evidence base changes. It will be impractical to resubmit RESPOND to a complete validation exercise as described in this article—nor is that necessary. Critical and essential will be to replicate Step 1: the verification by experts that an algorithm is an accurate and valid representation of the referent guideline.

A second challenge concerns the extent to which a guideline-based system like RESPOND reflects the complexity of clinical decision-making in the management of cancer-related anemia and the clinical dynamics of comorbidities in diagnosis, management, and prevention—and the clinical outcomes that can be achieved. Unfortunately, this is where the evidence-based healthcare model runs into a "self-imposed" limitation by attaching a higher grade of relevance to randomized controlled trials (and any ensuing meta-analyses) than to other research models. RCTs give the most reliable and internally valid information about the efficacy of a treatment (i.e., outcomes attributable to the intervention under investigation) by controlling the influence of extraneous variables through such sequential procedures as specifying inclusion/exclusion criteria, randomization to treatment arms, and *post hoc* statistical control of residual confounds. Typically, RCTs enroll relatively homogeneous groups of patients with limited variability in potential confounding variables. Comorbidities tend to be considered potential confounds and by and large are managed through exclusion—with only the more common ones being included. Consequently, comorbidities are less likely to be addressed in evidence-based guidelines; in contrast to the reality of daily clinical practice where patients are heterogeneous and treatment effectiveness may be confounded. The RESPOND system addresses this issue by allowing clinicians to override the guideline-recommended approach and documenting the reason for doing so.

Relatedly, a third challenge concerns the consideration of patient preferences in planning care. Patients choose between treatment options based on their personal evaluation of attributes, consequences, and the physical, psychological, and financials burdens involved. In RCTs, the building blocks of evidence-based guidelines, one implicitly assumes that patients have expressed their treatment preferences by providing informed consent to be included in the study. Actual patient preferences are not investigated in efficacy studies, even though later "real world" effectiveness may be affected by the extent to which patients agree to, accept, tolerate, and refuse the recommended treatment. As a guideline-driven

tool, RESPOND cannot accommodate patient preferences. However, being an advisory rather than a prescriptive tool, clinicians should feel comfortable overriding recommendations if these are contrary to patient preferences, and document the reasons why.

As the evidence-based guidelines movement is rapidly achieving a stage of maturity, in which major guidelines are developed for wide dissemination and to guide local guidelines, one can expect more systems like RESPOND to be developed. Important functional characteristics of such systems have already been proposed [17], however guidance as to the scientific validation of systems remains lacking. It is hoped that the methodology described here can be of assistance to other clinical areas. In the meanwhile, the RESPOND initiative, complex as it is, should provide the necessary evidence that the RESPOND system is based on algorithms that accurately specify the EORTC guidelines, and that it assists clinicians in caring safely for cancer patients with anemia in accordance with these guidelines.

Disclosures

Section I: Contributions to the project

- *Project concept*
  - *Methodological aspects and study paradigm*: I. Abraham, K. MacDonald, M. Turner, H. Warrinnier, R. Dunlop.
  - *Application to oncology*: M. Aapro, J. Van Erps, P. Soubeyran, H. Warrinnier, M. Turner, I. Abraham, K. MacDonald.
  - *Application to evidence-based medicine*: M. Aapro, P. Soubeyran, J. Van Erps, K. MacDonald, I. Abraham, M. Turner
- *Study design*: I. Abraham, M. Aapro, P. Soubeyran, J. Van Erps, M. Turner, K. MacDonald, M. Muenzberg.
- *Supervision of study implementation*: M. Turner, M. Muenzberg, K. MacDonald, I. Abraham.
- *Quality assurance*: M. Muenzberg, M. Turner, K. MacDonald, R. Dunlop, I. Abraham.
- *Statistical plan*: I. Abraham.
- *Manuscript preparation*: I. Abraham.
- *Critical review of manuscript for intellectual content*: J. Van Erps, M. Aapro, P. Soubeyran, K. MacDonald, R. Dunlop, M. Turner, M. Muenzberg, H. Warrinnier.

Section II: Conflict of interest

J. Van Erps has consulted with, received research grants and contracts from, and/or served as a sponsored speaker for the following companies: Roche, Novartis. She declares no conflict with regard to the work described in this manuscript.

M. Aapro has consulted with, received research grants and contracts from, and/or served as a sponsored speaker for the following companies: Roche, Amgen. He declares no conflict with regard to the work described in this manuscript.

P. Soubeyran has consulted with, received research grants and contracts from, and/or served as a sponsored speaker for the following companies: Roche, Amgen, Johnson & Johnson, Sanofi-Aventis, Schering AG, Schering-Plough, Pfizer, Chugai, and Baxter Oncology. He declares no conflict with regard to the work described in this manuscript.

R. Dunlop has consulted with, received research grants and contracts from, and/or served as a sponsored speaker for the following companies: Roche, Pfizer, Organon, NovoNordisk, and British Oxygen Company. InferMed has been contracted by Sponsor to provide support with software application development, testing, and support. He declares no conflict with regard to the work described in this manuscript.

I. Abraham and K. MacDonald have consulted with, received research grants and contracts from, and/or served as a sponsored speaker for the following companies and, as applicable, their subsidiaries: Novartis, Johnson & Johnson (including Centocor, Ortho-Biotech, Janssen Pharmaceutica, Janssen-Cilag, Janssen-Ortho), Eli Lilly, Roche, Pfizer, Amgen, Merck, Bristol-Myers Squibb, Schering-Plough, Astra-Zeneca, Bayer, GlaxoSmithKline, Lundbeck, and Innogenetics (including Xcellentis). Matrix45 has been contracted by Sponsor to provide support with project conceptualization, project design, protocol development, development of project materials, training, project management and implementation, development of statistical plan, and quality assurance. As per company policy, I. Abraham and K. MacDonald are barred from holding equity in any client companies and are subjected to internal and external review of their work to assure objectivity and transparency. They have taken the necessary steps to assure independence and do not declare a conflict of interest with regard to the work described in this manuscript.

M. Turner, M. Muenzberg, and H. Warrinnier, are employees of F. Hoffmann-La Roche and its subsidiaries. They have refrained from undue influence throughout the project and manuscript preparation.

Section III: Sponsor and role of sponsor

- *Sponsor*

  F. Hoffmann-La Roche AG (Basel, Switzerland) and its subsidiaries provide financial support for the project. Sponsor has also committed internal resources to support the project.
- *Role of the sponsor*

  The study paradigm was developed independently and proposed to sponsor by I. Abraham and K. MacDonald. The application to cancer-related anemia was identified by the sponsor. The independent oncology experts (J. Van Erps, M. Aapro, P. Soubeyran) were proposed by the sponsor. Collectively, all authors constituted the project team; employees of F. Hoffmann-La Roche AG listed as authors participated as members of the team. The manuscript was

developed by members of the research team not affiliated with Sponsor. Sponsor had right of review and comment. All final decisions regarding content of the manuscript were made by the external members of the team. See also Section I supra.

Section IV: Role of medical writer or editor

No medical writer or editor was retained in the preparation of this manuscript. M. Abraham provided proofreading support. K. Jones provided technical assistance in the preparation of the manuscript. Both were funded under the manuscript preparation subcontract by sponsor to Matrix45.

Reviewers

Jan Foubert, M.Sc., RPN, P.G. Dip., Ed. RNT, Departement Gezondheidszorg, Erasmushogeschool, Laarbeeklaan 121B-1090, Jette, Belgium.

Benjamin Djulbegovic, M.D., Ph.D., Department of Interdisciplinary Oncology, H. Lee Moffitt Cancer Center & Research Institute, USF, Tampa, FL 33612, United States.

Acknowledgments

Supported by research contracts from F. Hoffmann-La Roche AG. The authors thank the staff and patients at the center for their respective contributions to the study; an anonymous referee for most helpful suggestions; M. Abraham for independent proofreading support and K. Jones for technical assistance in manuscript preparation.

References

[1] Groopman JE, Itri LM. Chemotherapy-induced anemia in adults: incidence and treatment. J Natl Cancer Inst 1999;91:1616–34.
[2] Ludwig H, Van Belle S, Barrett-Lee P, et al. The European Cancer Anaemia Survey (ECAS): a large, multinational, prospective survey defining the prevalence, incidence, and treatment of anemia in cancer patients. Eur J Cancer 2004;40:2293–306.
[3] Weiss G, Goodnough LT. Anemia of chronic disease. New Engl J Med 2005;352:1011–23.
[4] Blohmer JU, Dunst J, Harrison L, et al. Cancer-related anemia: biological findings, clinical implications and impact on quality of life. Oncology 2005;68(Suppl. 1):12–21.
[5] Caro JJ, Salas M, Ward A, Goss G. Anemia as an independent prognostic factor for survival in patients with cancer: a systemic, quantitative review. Cancer 2001;91:2214–21.
[6] Nissenson AR, Goodnough LT, Dubois RW. Anemia: not just an innocent bystander? Arch Intern Med 2003;163:1400–4.
[7] Cella D. Factors influencing quality of life in cancer patients: anemia and fatigue. Semin Oncol 1998;25:43–6.
[8] Rizzo JD, Lichtin AE, Woolf SH, et al. Use of epoetin in patients with cancer: evidence-based clinical practice guidelines of the American Society of Clinical Oncology and the American Society of Hematology. J Clin Oncol 2002;20:4083–107.
[9] Bokemeyer C, Aapro MS, Courdi A, et al. EORTC guidelines for the use of erythropoietic proteins in anaemic patients with cancer. Eur J Cancer 2004;40:2201–16.
[10] Bokemeyer C, Aapro MS, Courdi A, et al. EORTC guidelines for the use of erythropoietic proteins in anaemic patients with cancer: 2006 update. Eur J Cancer 2007;43:258–70.
[11] Eddy DM. Evidence-based medicine: a unified approach. Health Aff 2005;24:9–17.
[12] Timmermans S, Mauck A. The promises and pitfalls of evidence-based medicine. Health Aff 2005;24:18–28.
[13] Halstenson CE, Macres M, Katz SA, et al. Comparative pharmacokinetics and pharmacodynamics of epoetin alfa and epoetin beta. Clin Pharmacol Ther 1991;50:702–12.
[14] Overbay DK, Manley HJ. Darbepoetin-alpha: a review of the literature. Pharmacotherapy 2002;22:889–97.
[15] Shiffman RN, Brandt CA, Liaw Y, Corb GJ. A design model for computer-based guideline implementation based on information management services. J Am Med Inform Assoc 1999;6:99–103.
[16] Sim I, Gorman P, Greenes RA, et al. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc 2001;8:527–34.
[17] Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003;10:523–30.
[18] Berner ES. Clinical decision-support systems: theory and practice. 2nd ed. New York: Springer-Verlag; 1998.
[19] Musen MA, Shahar Y, Shortliffe EH. Clinical decision support systems. In: Shortliffe EH, Cimino JJ, editors. Biomedical informatics: computer applications in health care and biomedicine. 3rd ed. New York: Springer Verlag; 2006.
[20] Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. Br Med J 2005;330:775–82.
[21] Nunnally JC, Bernstein I. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
[22] Park HM, Jung HW. Evaluating interrater agreement with intraclass correlation coefficient in SPICE-based software process assessment. In: IEEE proceedings, third international conference on quality software. Washington, DC: IEEE; 2003.
[23] Aapro M, Abraham I, MacDonald K, et al. Intraclass correlation metrics for the accuracy of algorithmic definitions in a computerized decision-support system for supportive cancer care. Support Care Cancer 2007;15:1325–9.

Biographies

Matti Aapro, M.D., is Dean of the Institut Multidisciplinaire d'Oncology, Clinique de Genolier, in Genolier, Switzerland. He was a member of the team that developed the EORTC guidelines for erythropoietic proteins in cancer-related anemia.

Joanna Van Erps, M.D., is head of the Departement Oncologie en Hematologie at Stedelijk Ziekenhuis Aalst, in Aalst, Belgium and a clinical preceptor, Faculty of Medicine, Vrije Universiteit Brussel, in Brussel (Belgium). She is the site investigator for the Step 2 validation study described in this article.

Karen MacDonald, Ph.D., is a Principal in Matrix45, a research and consulting firm specialized in phase 4 and 5 studies, evidence-based healthcare, and innovation in methods and statistics. Her main interests are in outcomes research

and modeling of processes and determinants of health and illness.

Pierre Soubeyran, M.D., Ph.D., is professor of medical oncology at the Université Victor Segalen Bordeaux 2 and head of clinical research at the Institut Bergonié, a comprehensive cancer center in Southwestern France. His research is focused on lymphomas and geriatric oncology. He was a member of the team that developed the EORTC guidelines for erythropoietic proteins in cancer-related anemia.

Matthew Turner, B.Sc. (Hon.), is associate medical director at Roche Products Limited, Welwyn Garden City, United Kingdom.

Michael Muenzberg, M.D., is International Medical Leader, Oncology, in the Anemia Business Franchise at F. Hoffmann-La Roche, Ltd. in Basel (Switzerland).

Robert Dunlop, M.D., is clinical director of InferMed Ltd., which specializes in the development of clinical decision support applications. His research is focused on extensions to the PROforma guideline representation language and the development of GRID-enabled clinical decision support.

Hans Warrinnier, M.D., is Medical Advisor, Anemia Nephrology/Oncology, at s.a. Roche n.v. in Brussel, Belgium.

Ivo Abraham, Ph.D. is a principal in Matrix45 and a professor at the University of Pennsylvania, where he is affiliated with the Center for Health Outcomes and Policy Research (School of Nursing), the Leonard Davis Institute of Health Economics (Wharton School of Business), and the Institute of Aging (School of Medicine). His research, both within Matrix45 and at the University of Pennsylvania, is focused mainly on practice patterns and outcomes in pharmacotherapeutics and medical devices across clinical areas, evidence-based healthcare and decision support, and the development and application of advanced research methods and statistics for health research.